DOCUMENT RESUME

ED 465 802 TM 034 189

AUTHOR Seltzer, Michael; Choi, Kilchan; Thum, Yeow Meng

TITLE Latent Variable Modeling in the Hierarchical Modeling Framework: Exploring Initial Status X Treatment Interactions in Longitudinal Studies. CSE Technical Report.

INSTITUTION National Center for Research on Evaluation, Standards, and Student Testing, Los Angeles, CA.; California Univ., Los Angeles. Center for the Study of Evaluation.

SPONS AGENCY Office of Educational Research and Improvement (ED), Washington, DC.

REPORT NO CSE-TR-559

PUB DATE 2002-03-00

NOTE 26p.

CONTRACT R305B60002

AVAILABLE FROM Center for the Study of Evaluation, National Center for Research on Evaluation, Standards, and Student Testing, Graduate School of Education & Information Studies, University of California at Los Angeles, 300 Charles E. Young Dr. North, Los Angeles, CA 90095-1522. Tel: 310-266-1532. For full text: http://www.cse.ucla.edu/CRESST/Reports/TECH559.PDF.

PUB TYPE Reports - Descriptive (141)

EDRS PRICE MF01/PC02 Plus Postage.

DESCRIPTORS *Bayesian Statistics; *Data Analysis; Estimation (Mathematics); Intervention

IDENTIFIERS *Hierarchical Linear Modeling; *Latent Variables; Time Series Analysis

ABSTRACT

In intervention studies, it is important to assess whether one program might be more effective for individuals with extreme initial difficulties, while another might be more effective for individuals with less extreme initial difficulties. In setting in which time-series data are obtained for each person, this entails examining interactions between treatment and initial status on rate of change. This report illustrates a fully Bayesian approach to studying interactions of this kind in the Hierarchical Modeling (HM) framework. This approach provides data analysis with a number of important advantages, including the ability to handle situations in which the number and spacing of time-series observations vary substantially across individuals and the ability to obtain robust estimates of parameters of interest. Various extensions of the approach are discussed in detail. (Contains 4 figures, 3 tables, and 22 references.) (Author/SLD)

CRESST
National Center for Research on Evaluation, Standards, and Student Testing

Latent Variable Modeling in the Hierarchical Modeling Framework: Exploring Initial Status x Treatment Interactions in Longitudinal Studies

CSE Technical Report 559

Michael Seltzer, Kilchan Choi, and Yeow Meng Thum
CRESST/University of California, Los Angeles

UCLA Center for the Study of Evaluation

In Collaboration With:

University of Colorado at Boulder • Stanford University • The RAND Corporation
University of Southern California • Educational Testing Service
University of Pittsburgh • University of Cambridge

Latent Variable Modeling in the Hierarchical Modeling Framework: Exploring Initial Status x Treatment Interactions in Longitudinal Studies

CSE Technical Report 559

Michael Seltzer, Kilchan Choi, and Yeow Meng Thum
CRESST/University of California, Los Angeles

March 2002

Center for the Study of Evaluation
National Center for Research on Evaluation,
Standards, and Student Testing
Graduate School of Education & Information Studies
University of California, Los Angeles
Los Angeles, CA 90095-1522
(310) 206-1532

Project 3.3 Validity of Measures of Progress. Bengt Muthén and Michael Seltzer, Project Directors, CRESST/GSEIS/University of California, Los Angeles.

The work reported herein was supported under the Educational Research and Development Centers Program, PR/Award Number R305B60002, as administered by the Office of Educational Research and Improvement, U.S. Department of Education.

The findings and opinions expressed in this report do not reflect the positions or policies of the National Institute on Student Achievement, Curriculum, and Assessment, the Office of Educational Research and Improvement, or the U.S. Department of Education.

Latent Variable Modeling in the Hierarchical Modeling Framework: Exploring Initial Status × Treatment Interactions in Longitudinal Studies

Michael Seltzer
CRESST/University of California, Los Angeles

Kilchan Choi
CRESST/University of California, Los Angeles

Yeow Meng Thum
CRESST/University of California, Los Angeles

Abstract

In intervention studies, it is important to assess whether one program might be more effective for individuals with extreme initial difficulties, while another might be more effective for individuals with less extreme initial difficulties. In settings where we obtain time-series data for each person, this entails examining interactions between treatment and initial status on rates of change. In this report, we illustrate a fully Bayesian approach to studying interactions of this kind in the Hierarchical Modeling (HM) framework. This approach provides data analysts with a number of important advantages, including the ability to handle situations in which the number and spacing of time-series observations varies substantially across individuals, and the ability to obtain robust estimates of parameters of interest. Various extensions of our approach are discussed in detail.

Many key questions in educational research, and in social and behavioral research more generally, entail measuring and studying change. In this connection, growth modeling techniques provide a valuable means of studying patterns of change over time (see, for example, Bryk & Raudenbush, 1987; Muthen, 1991).

Key substantive questions in studies of change often center on relationships between where individuals start (e.g., their initial status) and how rapidly they progress (e.g., their rates of change) (see Muthen & Curran, 1997; Khoo, 1997; Blomqvist, 1977). For example, in studying the effectiveness of two remedial reading programs over time, it becomes important to consider whether one program might be more effective for students with extreme reading difficulties, while the other program might

be more successful in the case of students with milder initial reading difficulties. Thus, among students with extreme reading difficulties, rates of progress may be more rapid for students in Program A, whereas among students with milder difficulties, rates of progress may be more rapid for students in Program B (see Figure 1). We term phenomena of this kind Initial Status × Treatment interactions.

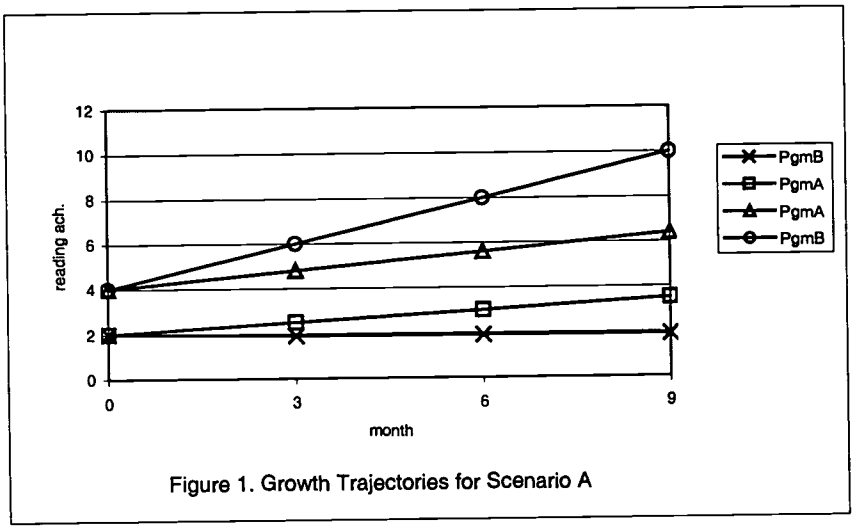


Figure 1. Growth Trajectories for Scenario A

As a second example, consider studying change in outcomes of interest in the context of large-scale educational surveys. In analyses of data from the National Longitudinal Survey of Youth (NLSY), for example, interest might center on how differences in initial status in anti-social behavior (*ASB*) relate to differences in rates of change in *ASB*: How much of a change in *ASB* growth rates do we expect when initial status increases 1 unit? Furthermore, interest might also center on how differences in various demographic characteristics (e.g., gender) and home environment factors, for example, correlate to differences in rates of change in *ASB*. It may well be the case that the effects of these factors depend crucially on (i.e., interact with) initial status.

Specifying models to address questions of this kind in essence implies modeling individual growth rate parameters (π_{1i}) as a function of individual initial status parameters (π_{0i}). Thus, for example, questions centering on Initial Status × Treatment interactions imply models for growth rate parameters of the following general form:

$$\pi_{1i} = f(\pi_{0i}, TRT_i, \pi_{0i} \times TRT_i) \qquad (1)$$

Viewing individual growth parameters as latent variables, the use of π_{0i} as a predictor of π_{1i} moves us into the arena of latent variable regression. While statistical analyses commonly are conducted using fallible measures of constructs of interest, a hallmark of Structural Equation Modeling (SEM) is that it provides a framework for specifying relationships among latent variables. As such, SEM provides an approach to growth modeling that enables us to employ initial status parameters as predictors of growth rates. SEM can be readily applied in settings in which we wish to specify Initial Status × Treatment interactions (e.g., $\pi_{0i} \times TRT_i$, where TRT_i is a 0/1 indicator variable), as well as situations in which interactions between π_{0i} and other predictors of growth are not specified (e.g., $\pi_{1i} = f(\pi_{0i}, TRT_i)$).

One limitation in current implementations of SEM is that our data must be time-structured. Thus, for example, the series of ages at which children in a developmental study are observed must be similar across children.

A second limitation that is less germane to the focus of this article is that it is not possible in standard implementations of SEM to specify interactions between initial status and continuous predictors of change. For example, consider an intervention study in which a key implementation variable is measured on a continuous scale. In settings of this kind, one would not be able to study interactions between initial status and implementation on rates of change.

In contrast to SEM, the hierarchical modeling (HM) framework can easily handle data sets that are not time-structured, i.e., settings in which the number and spacing of time series observations may vary across individuals. But unlike SEM, latent variable regression is not a hallmark of the HM framework.

In an important extension of the HM framework, Raudenbush and Sampson (1999) present a strategy for incorporating latent variable regressions into HMs. This strategy has been implemented in the latest release of the HLM software program. In its current implementation, this strategy enables one to employ initial status as a predictor of growth rates, but only in settings in which interactions between initial status and other predictors of change are not specified. A potentially valuable application of this approach would be intervention settings in which assignment to treatment and comparison groups is not random. One could, using this approach, study the effects of treatment on rates of change holding constant initial status (e.g., $\pi_{1i} = f(\pi_{0i}, TRT_i)$). Rather than adjusting with respect to observed outcome scores at time 1, which contain measurement

error, the adjustment would be based on a latent variable. Note that models of this kind can be fit readily via SEM.

In principle, the strategy presented by Raudenbush and Sampson could be extended to settings in which we wish to specify Initial Status × Treatment interactions. This would require the specification and estimation of separate variance-covariance matrices for the treatment and comparison groups. However, Raudenbush and Sampson's approach entails transforming ML estimates of variance components and fixed effects in standard HMs. As such, it appears that this strategy could not be extended to settings in which we wish to specify interactions between initial status and continuous predictors.

In fitting statistical models under normality assumptions, results are potentially vulnerable to outlying cases. Thus point estimates and standard errors for coefficients in latent regression models in both the SEM and HM frameworks may be non-robust to outlying individuals (e.g., a person whose rate of growth is unusually rapid). In addition, both the SEM framework and the strategy outlined by Raudenbush and Sampson, inferences are based on large-sample theory. Thus the use of these approaches in growth modeling settings in which the number of individuals is small or moderate may not be prudent.

In this paper, we present a fully Bayesian approach to latent variable regression in the HM framework. This approach entails calculating marginal posterior distributions of interest via a key Markov chain Monte Carlo (MCMC) technique, i.e., the Gibbs sampler. Our approach has the following strengths. First, it can be used to analyze data that are not time-structured. Secondly, it enables one to specify Initial Status × Treatment interactions. Third, it enables one to employ t distributional assumptions at any level of the HM, which has the effect of downweighting outliers. Note that the term outliers as used here refers both to outlying time-series observations (e.g., a time-series observation for an individual that is unusually high or low given the overall trend in that person's data), and to outlying individuals (e.g., a person whose rate of change is unusually rapid or slow in relation to other individuals). Fourth, inferences in the approach that we employ are not based on large-sample approximations. Fifth, rather than transforming estimates of fixed effects and variance components in standard HMs to obtain estimates of coefficients in latent variable regressions, our approach entails estimating these coefficients directly. This makes possible a number of useful modeling extensions which we discuss at the end of our paper (e.g., interactions between initial status and continuous predictors of growth).

Our fully Bayesian approach to latent variable modeling in the HM framework can be easily implemented using the software program BUGS (Spiegelhalter et al., 1996a), which is a near acronym for "Bayesian analysis Using the Gibbs Sampler". We illustrate our approach through analyses of the data from a randomized trial comparing two forms of short-term psychotherapy (see Svartberg, Seltzer & Stiles, 1998). The implications of our approach for the study of educational interventions and for constructing educational indicators are discussed at the end of our paper.

Illustrative Example

Background

From a pool of 20 individuals referred for short-term psychotherapy, 10 were randomly assigned to a directive, psychodynamic form of therapy termed STAPP, and 10 were randomly assigned to a non-directive form of therapy (NDP) (see Svartberg et al., 1998). A key outcome of interest in this study is level of client distress as measured by an instrument termed the Symptom Checklist-90 (SCL-90; Derogatis, 1977). Note that on the SCL-90 scale, scores between 0 and 0.20 indicate that an individual is asymptomatic; scores between 0.20-0.40 indicate mild levels of distress; scores between 0.40-1.00 indicate moderate levels of distress; and scores exceeding 1.00 indicate severe symptomology. Efforts were made to measure levels of distress at multiple points in time: immediately prior to the start of treatment, after 10 sessions, at termination, and 6, 12 and 24 months after termination. In our analyses, we focus on SCL-90 scores from the pre-intervention, 10-th session, and termination measurement occasions. See Svartberg et al. (1998) for a set of analyses that includes the post-intervention time points.

For both groups, treatment was to last for 20 sessions. However, for ethical reasons, treatment was prolonged in the case of one client for 32 sessions (client 9). This patient was assessed pretreatment, and at sessions 10, 20, and 32. In addition, one patient had only 14 sessions (client 2), and was assessed pretreatment and at sessions 10 and 14.

As in Svartberg et al. (1998), the carrier of time in our analyses of change is measured in units of months: $MONTH_{ti}$, where $MONTH_{ti}$ captures the number of months that have elapsed since the start of treatment for person i at measurement occasion t. Due to cancellations,

changes in schedules, patient and therapist vacations, and extensions of treatment in the case of 1 patient, $MONTH_{ti}$ takes on 57 different values ranging from 0.00 to 17.8 months, where 0.00 is the time-value corresponding to the pre-intervention measurement occasion. Thus the spacing between time-points varies considerably across patients. In addition, the duration of treatment ranges from 4.8 to 17.8 months. The average duration and median duration of treatment take on values of approximately 9 months and 8.8 months, respectively.

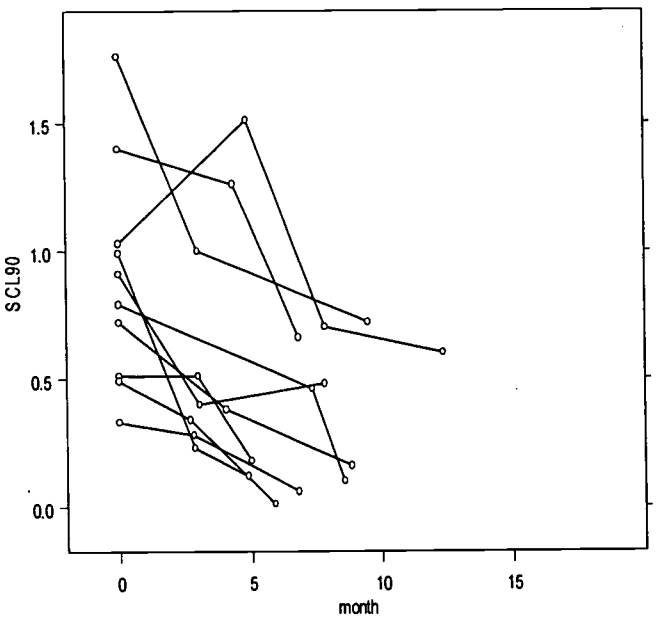


Figure 2A. Individual Growth Trajectories for StAPP Patients

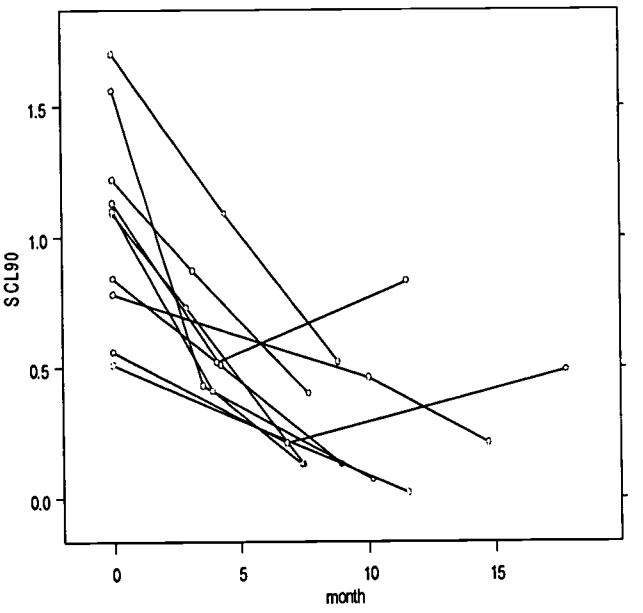


Figure 2B. Individual Growth Trajectories for NDP Patients

The trajectories of SCL-90 scores for the clients in the STAPP group and the clients in the NDP group are displayed in Figures 2a. and 2b., respectively. Note that at the outset, virtually all patients have SCL-90 scores that indicate moderate or high levels of distress. A key feature of these trajectories is that in general, SCL-90 scores tend to decrease over time in a fairly linear fashion. Three exceptions to this pattern are the trajectories for clients 9, 17 and 19.

We will first fit a growth model without latent variable predictors to the data in order to examine overall differences between STAPP and NDP patients in their initial status and rates of change. We will then employ initial status as a predictor of change and assess differences between STAPP and NDP patients in their rates of change holding constant initial status. Finally, we will examine whether the relative

effectiveness of STAPP and NDP on rate of change interacts with initial status.

Model I

We begin by specifying the following level-1 (or within-child) model:

$$Y_{ti} = \pi_{0i} + \pi_{1i} Month_{ti} + \varepsilon_{ti} \qquad \varepsilon_{ti} \sim N(0, \sigma^2), \qquad (2)$$

where Y_{ti} represents the SCL-90 score for individual i (i=1, ..., I) at measurement occasion t (t=1, ..., T_i), and $MONTH_{ti}$ captures the number of months that have elapsed since the start of treatment for person i at measurement occasion t. In this model, π_{0i} represents the SCL-90 status for person i at the start of treatment (i.e., initial status), and π_{1i} is the rate of change during treatment for person i. The ε_{ti} (i.e., the level-1 residuals) are assumed normally distributed with mean 0 and variance σ^2.

We now pose the following level-2 (or between-child) model:

$$\begin{aligned} \pi_{0i} &= \beta_{00} + \beta_{01} TRT_i + U_{0i} \qquad U_{0i} \sim N(0, \tau_{00}) \\ \pi_{1i} &= \beta_{10} + \beta_{11} TRT_i + U_{1i} \qquad U_{1i} \sim N(0, \tau_{11}), \end{aligned} \qquad (3)$$

where TRT_i = 0 if client i receives the STAPP treatment, and TRT_i=1 if client i receives the NDP treatment. By virtue of this coding scheme, β_{00} represents the expected initial status for STAPP patients, and β_{01} captures the overall difference in initial status between NDP and STAPP patients. Although random assignment was employed in this study, the number of patients in each group is small. Thus results for β_{01} provide us with a check on the comparability of patients in the two treatment groups. Turning to the level-2 equation for growth rates, β_{10} represents the expected rate of change in SCL-90 scores for STAPP clients, and β_{11} captures the overall difference in rates of change between NDP and STAPP patients. The U_{0i} and U_{1i} are level-2 residuals (i.e., random effects) assumed normally distributed with mean 0, and variance τ_{00} and τ_{11}, respectively. Note that τ_{00} captures the variance in initial status that remains after taking into account treatment group membership. Similarly, τ_{11} captures the residual variation in growth rates. Furthermore, $Cov(U_{0i}, U_{1i}) = \tau_{01}$, where τ_{01} is the covariance between initial status and rate of change for patients within each of the treatment groups.

Estimation

One widely used approach to estimation and inference for HMs is termed full maximum likelihood. This entails jointly estimating the fixed effects and variance components in an HM via maximum likelihood. Asymptotic standard errors are based on the Fisher information matrix. A second commonly used approach entails computing maximum likelihood estimates of the variance components in a given model, which is termed restricted maximum likelihood (REML) estimation. Generalized Least Squares is then used to obtain estimates and standard errors for the fixed effects. In this step, the variance-covariance parameters in the HM are set equal to their REML estimates.

In growth modeling settings in which the number of individuals is small, the above approaches can result in underestimates of uncertainty (e.g., standard errors that are too small), and point estimates that may constitute poor summaries of the data (see, e.g., Draper (1995), Rubin (1981) and Seltzer, Wong & Bryk (1996)). Note, however, that with respect to hypothesis tests for fixed effects, the HLM program employs critical values based on the family of t distributions. With the exception of settings in which datasets are highly unbalanced, this approach will tend to provide appropriate rejection rates.

The fully Bayesian approach entails basing inferences on the marginal posterior distributions of parameters of interest (e.g., $p(\beta_{11} | \mathbf{y})$). Such an approach involves specifying prior distributions for all unknowns in one's model. To obtain the marginal posterior distribution of a particular parameter, one must integrate over all other unknowns in one's model. Thus, for example, $p(\beta_{11} | \mathbf{y})$ would provide us with a summary of the plausibility of different values for $p(\beta_{11} | \mathbf{y})$ given the data at hand and any available prior information. The mode, median and mean of $p(\beta_{11} | \mathbf{y})$ would provide us with various point estimates for β_{11} and the .025 and .975 quantiles of this distribution would provide us with the Bayesian analogue of a confidence interval.

An advantage of the fully Bayesian approach is that it provides a general strategy for drawing inferences concerning a parameter of interest in a manner that takes into account the uncertainty connected with all other unknowns in one's model. For example, in drawing inferences concerning β_{11}, integrating over the variance components as well as all other unknowns in effect propagates the uncertainty concerning these parameters into $p(\beta_{11} | \mathbf{y})$ (see, e.g., Draper (1995), Rubin, 1981, and Seltzer et al. (1996)).

Calculating marginal posteriors of interest has until recently been intractable in all but the simplest HM settings. However, MCMC techniques (e.g., the Gibbs sampler) now make such an approach feasible in a wide range of complex modeling settings (see, e.g., Carlin & Louis, 1996; Gelfand, Hills, Racine-Poon & Smith, 1990; Gelman, Carlin, Stern & Rubin, 1995; Seltzer et al., 1996; Spiegelhalter et al., 1996b, 1996c; Tanner, 1996). As will be seen, MCMC can be used to obtain marginal posteriors of interest in HMs in which level-1 parameters (e.g., π_{0i}) are employed as predictors at level-2, and in which t distributional assumptions are employed at any level of the model. Rather than transforming ML estimates of variance components and fixed effects in standard HMs to obtain estimates of coefficients in latent variable regressions, a distinct advantage in using MCMC is that these coefficients can be estimated directly. This makes possible a number of important extensions of our approach, which will be discussed at the end of our paper.

We used the software package BUGS to carry out all of the fully Bayesian analyses presented in this paper. BUGS, which is freely available via the Web, provides a relatively easy means of implementing the Gibbs sampler in a wide array of modeling settings. We ran the BUGS package on a Pentium II 400mhz PC. For each analysis, less than 15 seconds of CPU time were required to complete 10,000 iterations of the Gibbs sampler.

To diagnose possible convergence problems, for each analysis we ran multiple chains of the Gibbs sampler using different starting values and seeds, compared results based on the output from each chain, inspected trace plots, and examined Raftery-Lewis statistics. These procedures failed to identify any convergence problems. Note that trace plots, Raftery-Lewis statistics and a number of other useful convergence diagnostics can be obtained using a suite of programs called CODA (Best, Cowles & Vines, 1995), which has been made available by the developers of BUGS. Like BUGS, CODA is also freely available via the Web. To help ensure results with high degrees of accuracy, we employed a burn-in period of 2,000 iterations, and used the output from 60,000 subsequent iterations of the Gibbs sampler to simulate marginal posteriors of interest.

We specified diffuse priors for the fixed effects in our models. For example, in the case of the fixed effects in Model I, we employed independent normal priors with means of 0 and extremely low precision. Note that in BUGS, one routinely works with precisions (e.g., $1/\sigma_2$, $\mathbf{T}^{-1}$) rather than variances (σ_2, $\mathbf{T}$). The approach that we used to specify priors for precision parameters parallels the approach employed in papers by

Seltzer et al. (1996), Seltzer, Novak, Lim and Choi (2001), and Seltzer and Choi (in press) for specifying diffuse priors for variance components.

Results for Model I

As can be seen in Table I, the marginal posterior mean of β_{00} takes on a value of 0.87, which falls just below the lower boundary of the high distress category. The mean of the resulting marginal posterior distribution for β_{01} (i.e., the expected difference in initial status between NDP and STAPP patients) is slightly under a tenth of a point, and the 95% interval based on this distribution comfortably includes a value of 0. Thus, on average, the NDP and STAPP patients appear to be fairly similar in terms of initial levels of distress.

Table I: Marginal Posterior Distributions for the Fixed Effects and Variance Components in Model I.

	Mean	SD	95% Int.	Median	Prop.>0
Fixed Effects :					
Model for Initial Status (π_{0i})					
STAPP (β_{00})	.874	.138	(.596,1.145)	.875	1.000
NDP/STAPP Contrast (β_{01})	.088	.197	(-.305, .477)	.090	.681
Model for Rates of Change (π_{1i})					
STAPP (β_{10})	-.070	.020	(-.109,-.032)	-.070	.000
NDP/STAPP Contrast (β_{11})	-.002	.026	(-.053, .049)	-.002	.473
Variance Components :					
Within-Person Error (σ^2)	.060	.017	(.035, .100)	.057	
Random Effects Variance for Initial Status (τ_{00})	.149	.065	(.063, .308)	.136	
Random Effects Variance for Rates of Change (τ_{11})	.002	.001	(.001, .004)	.002	
Cov. between Init. Status and Rates of Chagne(τ_{01})	-.007	.006	(-.022, .002)	-.006	.074
Correlation between Initial Status and Rates of Change (τ_{01} / [$\tau_{00}^{1/2}$ x $\tau_{11}^{1/2}$])	-.407	.253	(-.797, .173)	-.443	.074

Turning to the results for the fixed effects in the level-2 equation for growth rates, we see that marginal posterior mean for β_{10} takes on a value of -0.07, and that the 95% marginal posterior interval excludes a value of 0. The value -0.07 suggests that on average, we expect to see a 0.07 decrease in a patient's SCL-90 score for each month of treatment that elapses. The results for β_{11} suggest a negligible difference in rates of improvement between NDP and STAPP patients. Note that the posterior mean for β_{11} is close to a value of 0, and that the value $\beta_{11} = 0$ falls near the center of the 95% marginal posterior interval.

In terms of results for the variance-covariance components in the model, the results for τ_{01} are of particular interest. Note that approximately 92% of the mass of the marginal posterior distribution for τ_{01} lies below a value of 0, thus providing some evidence (albeit modest evidence) of a negative relationship between initial status and rate of change (e.g., clients with high initial status values tend to exhibit more rapid rates of decline in their SCL-90 scores). (Note that the column in Tables I, II and III labelled "Prop. > 0" denotes the proportion of values among the set of 60,000 values generated for a parameter that exceed a value of 0. These proportions constitute highly accurate estimates of posterior probabilities.) In using MCMC techniques, one can also readily obtain the marginal posterior distributions of parameters that are combinations of other parameters in one's model. Thus using values generated in M cycles of the Gibbs sampler to simulate marginal posteriors of interest, and setting M to a large value, the empirical distribution of the values $\tau_{10}^{(i)}/[\tau_{00}^{1/2(i)}\tau_{11}^{1/2(i)}]$ (i=1,...., M) provides us with an accurate approximation of the marginal posterior distribution of the correlation between initial status and rate of change for patients within the two treatment groups. As can be seen in Table I, the mean of the resulting posterior takes on a value of approximately -0.41, and over 92% of the posterior lies below a value of 0.

Model II

Though the results for Model I point to very little difference in initial status between STAPP and NDP clients, we now employ initial status (π_{0i}) as a latent variable predictor in the level-2 equation for growth rates. This will enable us to illustrate a potentially important use of initial status as a latent variable predictor of change in intervention settings, i.e., assessing the relative effectiveness of interventions on rates of change holding constant initial status.

We now pose the following latent variable regression HM. We employ the same level-1 model as in Equation 2. At level-2, we specify the following model:

$$\pi_{0i} = \beta_{00} + U_{0i} \qquad U_{0i} \sim N(0, \tau_{00})$$
$$\pi_{1i} = \beta_{10} + \beta_{11} TRT_i + b\pi_{0i} + U_{1i} \qquad U_{1i} \sim N(0, \tau_{11}). \qquad (4)$$

In the first equation in the level-2 model, the individual initial status parameters are modeled as a function of a grand mean (i.e., β_{00}). The random effects in the first equation now represent deviations in initial status from the grand mean (i.e., $U_{0i} = \pi_{0i} - \beta_{00}$), and τ_{00} captures the variance in the π_{0i} around the grand mean.

Note importantly that in contrast to Model I, π_{0i} now appears as a predictor in the level-2 equation for rates of change. In this model, β_{11} represents the expected difference in rates of change between NDP and STAPP patients holding constant initial status. The parameter b is the regression coefficient for the latent predictor π_{0i}. b is a pooled slope for patients in the NDP and STAPP treatment groups that captures the expected change in growth rate when initial status increases one unit.

The parameter β_{10} represents the expected rate of change for a STAPP patient (i.e., $TRT_i = 0$) with an initial status value of 0. To give β_{10} a more meaningful interpretation, we can center π_{0i} around the grand mean for initial status as follows:

$$\pi_{1i} = \beta_{10} + \beta_{11} TRT_i + b(\pi_{0i} - \beta_{00}) + U_{1i} \qquad U_{1i} \sim N(0, \tau_{11}). \qquad (5)$$

In this model, β_{10} now represents the expected rate of change for a STAPP patient whose initial status value is equal to the grand mean.

The variance parameter in the level-2 equation for growth rates (i.e., τ_{11}) represents the amount of variance in growth rates that remains after taking into account initial status and treatment group membership. Since we are conditioning on initial status in the level-2 model for growth rates (i.e., $\pi_{1i} \mid \pi_{0i}, TRT_i$), we assume that $Cov(U_{0i}, U_{1i}) = 0$. Assumptions of this kind are made routinely in the SEM growth modeling framework in settings where initial status is employed as a predictor of rates of change.

Table II: Marginal Posterior Distributions for Fixed Effects and Variance Components in Model II.

	Mean	SD	95% Int.	Median	Prop.>0
Fixed Effects :					
Model for Initial Status (π_{0i})					
Grand Mean (β_{00})	.913	.100	(.715, 1.108)	.912	1.000
Model for Rates of Change (π_{1i})					
STAPP (β_{10})	-.071	.018	(-.107,-.036)	-.071	.999
NDP/STAPP Contrast (β_{11})	.004	.020	(-.035, .043)	.004	.572
Init. Status Effect(b)	-.067	.035	(-.130, .007)	-.069	.034
Variance Components :					
Within-Person Error (σ^2)	.061	.017	(.035, .102)	.058	
Random Effects Variance for Initial Status (τ_{00})	.149	.067	(.064, .307)	.137	
Random Effects Variance for Rates of Change (τ_{11})	.0010	.0006	(.0003,.0024)	.0009	

As can be seen in Table II, the resulting posterior mean for the NDP/STAPP difference in growth rates holding constant initial status is extremely close to a value of 0. We also see that the marginal posterior mean for b takes on a value of -0.067, which implies that a 1 unit increase in initial status implies a decrease in rate of change in SCL-90 scores (i.e., a more rapid decrease in distress) of 0.067 units per month. While a value of 0 lies within the upper boundary of the 95% interval for b, note that only approximately 3.4% of the mass of the posterior lies above a value of 0. Results for β_{10} imply an expected rate of change of -0.071 for a STAPP patient with an initial status value equal to the grand mean.

Model III

The above analyses point to the conclusion that the effects of STAPP and NDP on rates of change in SCL-90 scores do not differ. However, when we examine the observed SCL-90 trajectories for STAPP and NDP patients, we see that rates of change among the STAPP patients tend to be fairly similar regardless of where patients start, i.e., irrespective

of their pre-intervention SCL-90 scores (see Figure 2a). In contrast, in Figure 2b we see that NDP patients with high pre-intervention scores tend to exhibit rapid decreases in distress, while those with relatively lower pre-intervention scores tend to progress at substantially slower rates. This suggests that NDP may be more effective for patients with high initial levels of distress while STAPP may be more effective for patients with moderate initial levels of distress.

To investigate this possibility we expand Equation 5 to include an Initial status × Treatment interaction term:

$$\begin{aligned} \pi_{0i} &= \beta_{00} + U_{0i} \qquad\qquad U_{0i} \sim N(0, \tau_{00}) \\ \pi_{1i} &= \beta_{10} + \beta_{11} TRT_i + b_1(\pi_{0i} - \beta_{00}) + b_2 [TRT_i \times (\pi_{0i} - \beta_{00})] + U_{1i} \\ &\qquad\qquad U_{1i} \sim N(0, \tau_{11}), \end{aligned} \tag{6}$$

where β_{11} and b_1 represent, respectively, the main effects of treatment and initial status on rates of change, and where b_2, the parameter of primary interest, captures the interaction between initial status and treatment on rates of change.

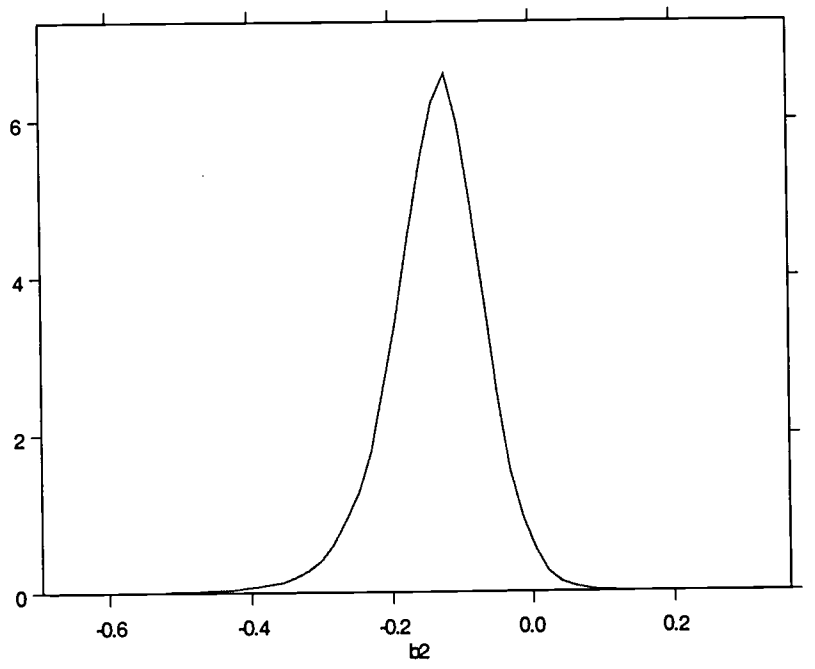


Figure 3. The Marginal Posterior Distribution of the Initial × Treatment Interaction Effect (b_2)

Based on the equation for growth rates in the above level-2 model, it can be seen that the expected rate of change for STAPP patients ($TRT_i = 0$) is as follows:

$$E(\pi_{1i} \mid TRT_i = 0) = \beta_{10} + b_1(\pi_{0i} - \beta_{00}). \quad (7)$$

For NDP patients, the expected rate of change is:

$$E(\pi_{1i} \mid TRT_i = 1) = (\beta_{10} + \beta_{11}) + (b_1 + b_2)(\pi_{0i} - \beta_{00}). \quad (8)$$

In Table III, we see that the mean of the resulting posterior distribution for b_2 takes on a value of -0.137 and that the 95% interval for b_2 includes only negative values (see also Figure 3). Note also that approximately 98% of the mass of $p(b_2 \mid y)$ lies below a value of 0. These results point strongly to an interaction between initial status and treatment on rate of change.

Table III: Marginal Posterior Distributions for the Fixed Effects and Variance Components in Model III

	Mean	SD	95% Int.	Median	Prop.>0
Fixed Effects :					
Model for Initial Status (π_{0I})					
Grand Mean (β_{00})	.923	.098	(.728, 1.111)	.924	1.000
Model for Rates of Change (π_{1i})					
STAPP (β_{10})	-.073	.017	(-.106,-.041)	-.073	.000
NDP/STAPP Contrast (β_{11})	.004	.022	(-.039, .048)	.004	.570
Initial Status Effect(b1)	.003	.053	(-.085, .124)	-.002	.479
Init. Status x Treatment (b2)	-.137	.071	(-.291,-.012)	-.132	.016
Variance Components :					
Within-Person Error (σ^2)	.057	.015	(.034, .094)	.055	
Random Effects Variance for Initial Status (τ_{00})	.141	.060	(.060, .290)	.129	
Random Effects Variance for Rates of Change (τ_{11})	.0008	.0004	(.0002, .0019)	.0007	

To help interpret the results of this analysis, we use the resulting posterior means of the fixed effects to compute expected rates of change for STAPP patients and for NDP patients whose initial status values are one standard deviation above or below the grand mean. First, since τ_{00} represents the variance in initial status values around the grand mean, then $(\pi_{0i} - \beta_{00}) = \tau_{00}^{1/2}$ represents the grand-mean centered initial status value for a person whose initial status value is one standard deviation above the mean. Similarly, $(\pi_{0i} - \beta_{00}) = -\tau_{00}^{1/2}$ represents the centered initial status value for a person whose initial status value is one standard deviation below the grand mean. Since the posterior distribution for τ_{00} is somewhat skewed (see Table III), the median of this distribution (0.129) provides us with a sensible estimate of τ_{00} for the purpose at hand.

For STAPP patients whose initial status is one standard deviation above average ($(\pi_{0i} - \beta_{00}) = (0.129)^{1/2} = 0.36$), the expected rate of change based on our fitted model is:

$$E(\pi_{1i} \mid TRT_i = 0, [\pi_{0i} - \beta_{00}] = 0.36) = -0.073 + 0.003(0.36) = -0.072. \quad (9)$$

where -0.073 and 0.003 are the marginal posterior means of β_{10} and b_1, respectively.

For NDP patients whose initial status is one standard deviation above average, the expected rate of change is:

$$E(\pi_{1i} \mid TRT_i = 1, [\pi_{0i} - \beta_{00}] = 0.36) = (-0.073 + 0.004) + (0.003 + [-0.137])(0.36) = -0.117, \quad (10)$$

where -0.073, 0.004, 0.003 and -0.137 are the marginal posterior means of β_{10}, β_{11}, b_1 and b_2, respectively.

Thus among patients whose initial status value is one standard deviation above the grand mean, NDP patients are expected to improve at an appreciably more rapid rate than STAPP patients. In Figure 4, we see that after 6 months of treatment, SCL-90 scores are expected to be approximately .27 points lower for NDP patients than for STAPP patients.

For STAPP patients whose initial status is one standard deviation below average ($(\pi_{0i} - \beta_{00}) = (0.129)^{1/2} = -0.36$), the expected rate of change based on our fitted model is:

$$E(\pi_{1i} \mid TRT_i = 0, [\pi_{0i} - \beta_{00}] = -0.36) = -0.073 + 0.003(-0.36)$$

$$= -0.074. \quad (11)$$

For NDP patients whose initial status is one standard deviation below average, the expected rate of change is:

$$E(\pi_{1I} \mid TRT_i = 1, [\pi_{0I} - \beta_{00}] = -0.36) = (-0.073 + 0.004) + (0.003 + [-0.137])(-0.36)$$
$$= -0.117. \quad (12)$$

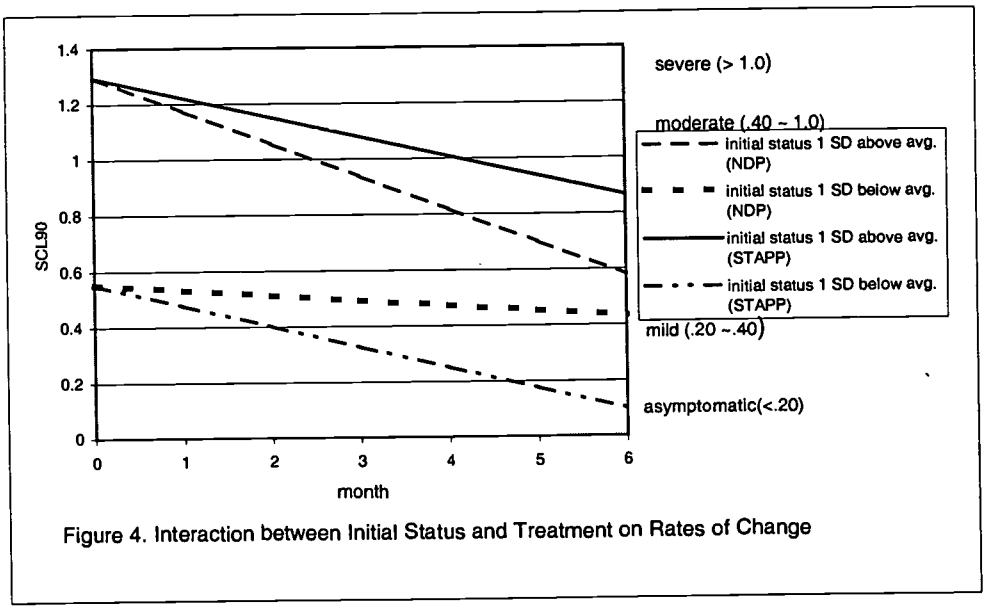


Figure 4. Interaction between Initial Status and Treatment on Rates of Change

As can be seen, the expected rate of change for low initial status STAPP patients (-0.074) is extremely similar to the expected rate for high initial status STAPP (-0.072) patients. Moreover, among low initial status patients, STAPP patients are expected to improve at substantially more rapid rates than NDP patients (-0.074 versus -0.021). As can be see in Figure 4, the expected rate of change for low initial status NDP patients is quite slow. After 6 months of treatment, SCL-90 scores for low initial status patients are expected to be approximately 0.30 points lower for STAPP patients than NDP patients.

To study the sensitivity of our results to possible outlying time-series observations or outlying individuals, we re-analyze the data employing heavy-tailed distributional assumptions at levels 1 and 2 (i.e., t distributional assumptions with 4 degrees of freedom): $\varepsilon_{ti} \sim t(0, \sigma^2, 4)$; $U_{0i} \sim t(0, \tau_{00}, 4)$; and $U_{1i} \sim t(0, \tau_{11}, 4)$. The results that we obtain are extremely similar to those obtained under normality assumptions. The resulting

marginal posterior mean for the interaction effect (b_2) takes on a value of -0.132, and we obtain a 95% interval for b_2 that includes only negative values (i.e., -0.301, -0.006).

While Models I and II point to virtually no difference in the effectiveness of STAPP and NDP, Model III points to NDP being more effective for patients whose initial status values are relatively high and to STAPP being more effective in the case of patients whose initial status values are relatively low.

Discussion

An important implication of the above analyses is that the ability to fit models that contain Initial Status × Treatment interaction terms encourages us to search for potentially important interactions that might otherwise go unnoticed. Our model-fitting approach is based on the use of MCMC techniques, which provide a viable means of obtaining estimates and standard errors for parameters of interest in numerous complex modeling settings. A key advantage of utilizing MCMC is that we are able to estimate coefficients in latent variable regressions directly as opposed transforming ML estimates of variance components and fixed effects in standard HMs. Various possible modeling extensions based on this estimation approach are as follows.

1. As noted above, our MCMC-based approach makes it possible to specify interactions between initial status and continuous predictors in modeling rates of change. This, in turn, potentially broadens the kinds of questions that we are able to address via analyses of the data from longitudinal educational surveys (e.g., NELS, ECLS), and from longitudinal studies of educational programs. For example, consider an intensive longitudinal study of an innovative remedial reading intervention, and that a key feature of the study is that children are randomly assigned to various versions of the program that differ in intensity i.e., in the number of minutes of remedial instruction per week). Thus, of particular interest are the effects of the intensity of treatment (*INTENSITY*) on rates of change, and the issue of whether, for example, differences in intensity are more consequential for those children who are most in need of the intervention (i.e., those children with markedly low initial status values). This implies a model for rates of change such as the following:

$$\pi_{1i} = \beta_{10} + \beta_{11} INTENSITY_i + b_1(\pi_{0i} - \beta_{00}) + b_2 [INTENSITY_i \times (\pi_{0i} - \beta_{00})] + U_{1i} \qquad U_{1i} \sim N(0, \tau_{11}), \qquad (13)$$

In this model, the key parameters of interest (i.e., b_1 and b_2) can be estimated directly via MCMC.

2. Student progress of course occurs in school settings. Building on Burstein's work on multilevel analysis (1980), relationships of substantive importance (e.g., the relationship between initial status and rate of progress) can vary substantially across key organizational units (e.g., classrooms, schools). For example, in some schools, where a student starts with respect to reading achievement may be extremely consequential in terms of how rapidly he or she progresses (e.g., those students who start high might progress rapidly, while those who start low might progress very slowly). In other schools, however, where a student starts may be less consequential in terms of how quickly he or she progresses. For example, in some schools, all children, regardless of their initial status, may progress rapidly. It therefore becomes important to examine how differences in various school policies and practices (e.g., differences in instructional materials, in the allocation of instructional time, in types and amounts of remedial services) affect the relationship between where students start and their subsequent rates of progress. This can be accomplished through the use of three-level HMs, which can be estimated via our MCMC approach.

3. Our MCMC-estimation strategy can be used in studies in which primary interest centers on the relationship between growth in different domains. For example, one can specify and fit models in which rate of change in word recognition skills is used as a predictor of subsequent rate of change in reading comprehension.

Our work in this paper has implications for the development of educational indicators. In particular, examining relationships between where students start and how rapidly they progress, and examining interactions involving various student background characteristics (e.g., SES, gender), help provide a sense of the kinds of students in a school who are benefiting and those who are not. This is the focus of a forthcoming deliverable.

References

Best, N., Cowles, M. & Vines, K. (1995). CODA: Convergence diagnosis and output analysis software for Gibbs sampling output,Version 0.30. MRC Biostatistics Unit, Cambridge University.

Blomqvist, N. (1977). On the relation between change and initial value. *Journal of the American Statistical Association*, 72, 746-749.

Bryk, A. S. & Raudenbush, S. W. (1987). Application of hierarchical linear models to assessing change. *Psychological Bulletin*, 101, 147-158.

Burstein, L. (1980). The analysis of multilevel data in educational research and evaluation. *Review of Research in Education*, 8, 158-233.

Carlin, B. & Louis, T. (1996). *Bayes and empirical Bayes methods for data analysis*. London: Chapman and Hall.

Derogatis, L. (1977). *SCL-90 manula-1*. Baltimore: Clinical Psychometrics Research Unit, Johns Hopkins University School of Medicine.

Draper, D. (1995). Inference and hierarchical modeling in the social sciences. *Journal of Educational and Behavioral Statistics, 20*, 115-147.

Gelfand, A., Hills, S., Racine-Poon, A. & Smith, A. (1990). Illustration of Bayesian inference in normal data models using Gibbs sampling. *Journal of the American Statistical Association, 85*, 972-985.

Gelman, A., Carlin, J., Stern, H. & Rubin, D. (1995). *Bayesian data analysis*. London: Chapman & Hall.

Khoo, S. (1997). *Assessing interactions between program effects and baseline: A latent curve approach*. Unpublished doctoral dissertation: University of California, Los Angeles.

Muthen, B. (1991). Analysis of longitudinal data using latent variable models with varying parameters. In L. Collins & J. Horn (Eds.), *Best methods for the analysis of change: Recent advances, unanswered questions, future directions* (pp. 1-17). Washington D.C.: American Psychology Association.

Muthen, B. & Curran, P. (1997). General longitudinal modeling of individual differences in experimental designs: A latent variable framework for analysis and power. *Psychological Methods*, 2, 371-402.

Raudenbush, S. & Sampson, R. (1999). Assessing direct and indirect associations in multilevel designs with latent variables. *Sociological Methods and Research*, 28(2),123-153.

Rubin, D. B. (1981). Estimation in parallel randomized experiments. *Journal of Educational Statistics*, 6, 377-400.

Seltzer, M. & Choi, K. (2000). Sensitivity analysis for multilevel models. In N. Duan & S. Reise (Eds.), *Multilevel modeling: Methodological advances, issues, and applications*. Mahwah, NJ: Lawrence Earlbaum.

Seltzer, M., Novak, J., Choi, K. Lim, N. (2001). Sensitivity for hierarchical models employing *t* level-1 assumptions. *Journal of Educational and Behavioral Statistics*. Manuscript conditionally accepted for publication.

Seltzer, M., Wong, W.H. & Bryk, A. (1996). Bayesian inference in applications of hierarchical models: Issues and methods. *Journal of Educational and Behavioral Statistics, 21*, 131-167.

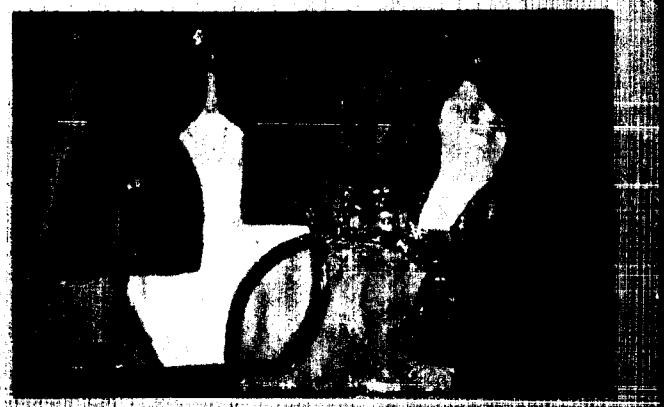
Spiegelhalter, D., Thomas, A., Best, N. & Gilks, W. (1996a). *BUGS: Bayesian inference using Gibbs sampling, Version 0.50*. MRC Biostatistics Unit, Cambridge University.

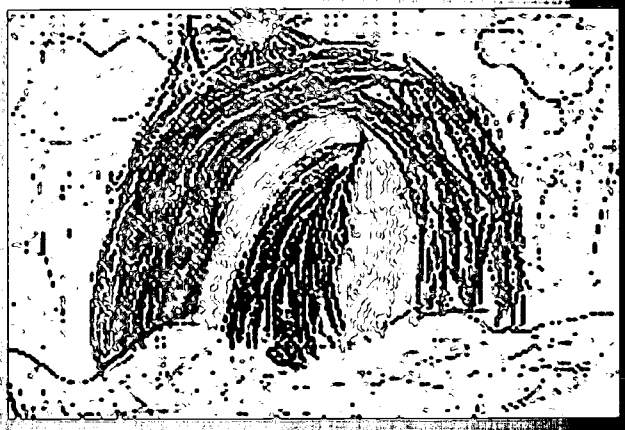
Spiegelhalter, D., Thomas, A., Best, N. & Gilks, W. (1996b). *BUGS examples volume 1, Version 0.50*. MRC Biostatistics Unit, Cambridge University.

Spiegelhalter, D., Thomas, A., Best, N. & Gilks, W. (1996c). *BUGS examples volume 2, Version 0.50*. MRC Biostatistics Unit, Cambridge University.

Svartberg, M., Seltzer, M. & Stiles, T. (1998). The effects of common and specific factors in short-term anxiety-provoking psychotherapy. *Journal of Nervous and Mental Disease*, 186, 691-696.

Tanner, M. (1996). *Tools for statistical inference* (3rd ed.). New York: Springer-Verlag.






Auriel

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

TM034189

NOTICE

Reproduction Basis

☒ This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

☐ This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").